Universitat de Lleida

TREBALL FINAL DE GRAU

ESCOLA POLITÈCNICA SUPERIOR
UNIVERSITAT DE LLEIDA
INSPIRING THE FUTURE

Estudiant: Iván Méndez Lobato

Titulació: Grau en tècniques d'interacció digital i de computació

Títol de Treball Final de Grau: **SEO Analysis and its effects on Web Positioning**

Director/a: **Jordi Mateo**

Presentació

Mes: Setembre

Any: 2021

SEO Analysis and its effects on Web Positioning

Iván Méndez, Jordi Mateo-Fornes

Abstract—Web and Web Applications are growing strong within the foreseeable future. In this context, and considering the non-stop growth of the online market, it is more important than ever to have a favorable positioning in search engines. Knowing the true nature of the positioning algorithm is a big challenge, as it is usually intentionally obscured by the Search Engine Companies, commonly hiding behind a Good Practices document or general guidelines. This study aims to research the inner workings of SEO and improve our organic position in the Google Search Engine ranking. We are dissecting and analyzing the features and characteristics more desirable or impactful in a well-positioned Web. The method for our study will revolve about starting with a basic, unoptimized Web, and introducing, ever so slightly, modifications and upgrades to our site, measuring how its SEO ranking grew over time and how much—listing and explaining not only the good practices to follow but also detailing practices, structures and strategies to avoid and why. The results obtained provide helpful, straight-to-the-point guidelines and recommendations with measurable results for Web developers and site owners.

Index Terms—SEO, Google Search Engine Positioning

1 Introduction

The Internet has evolved massively over time, and it will keep doing just that. Starting with the first proposal for the World Wide Web in March 1989 and a second, more polished proposal in 1990 [1], the outline for the Web's main concepts and ideas came to be. Since then, the popularity of the Web has grown tremendously. More than half of the world population has access to the Internet. With new ways of satellite broadband Internet systems like Starlink [2], it is easy to see that this percentage will only grow over time.

Nowadays, over 27000 new users access for the first time, every hour. Considering that these users come primarily from developed countries or countries currently developed, the new advances of fiber-less Internet access exacerbate the importance of online presence for marketing and selling purposes. Web usage has grown, but the capabilities and features of the Web itself have been evolving too. Starting with the read-only **Web 1.0** with simple features like static pages and email only forms [3], **Web 2.0** started spreading after 2002, known as the *wisdom Web* offered new features and services such as blogs, tags, simple RSS [4], writing technologies and wikis. After that, the Web evolved to the **Web 3.0** or the *semantic Web*. Its purpose was to make webs readable not only by humans but also by machines. In the last place, we get to the **Web 4.0**, an idea in progress, the *symbiotic Web* focused on creating relationships between men and machines.

Along with the boom in the number of websites, the need for a tool to accurately find the desired information was in dire need. Web Search engines were the answers to meeting these needs. Search engines work by discovering and cataloging the reachable content by crawling, indexing, and ordering the results by matching them to the searched words or query [5].

- *Iván Méndez Lobato, Jordi Mateo-Fornes (corresponding author) are with the Departament of Computer Science & INSPIRES, University of Lleida, Av. Jaume II 69, 25001 LLeida, Spain. Emails: iml1@alumnes.udl.cat, jordi.mateo@udl.cat*

Starting with rudimentary FTP crawlers [6], and evolving to keyword searching technologies, presented a way to search for manually or automatically generated indexes. It was not until 1994 that the first web crawlers allowed the users to search for any word in any part of the document. It was then seen, again, as an advertising opportunity for businesses.

Focusing on selling products and online services, and also following the rising trend of the Web, e-commerce emerged, and shortly after, the marketplaces were not limited to an exclusive seller, previously known as the owner of the site, but favored selling products by multiple third parties, connecting buyers and sellers. The most prominent marketplace, Amazon, sells more than 400 items a minute [7]. Retail opted first to keep divided its online and physical selling. Still, it soon realized the benefits of integrating both sides of the business, one of them being the possibility of adopting multichannel marketing practices [8].

One of those practices being SEO [9] (Search Engine Optimization) and SEM [10] (Search Engine Marketing). SEO centers on the organic results, in other words, the non-paid results presented in our Search Engine. It consists of a series of techniques applied to our Web to obtaining better positioning. It is composed of two main factors: Relevance and Authority.

- The **relevance** of a site is the measure in which a site is an exact match for a detailed query. This measure is the fundamental aspect we can apply to the majority of the techniques we will see later.
- The **authority** refers to the popularity of the site, the main indicator of it, is the number of out-site links to our Web

SEM consists of techniques to improve our site positioning employing paid advertisement in the Search Engines, i.e., Google Ads [11]. At its core, it is a trade-off between showing the site in the first set of results or in a privileged section of the results and the user of the SEM strategy pay-

ing for every click a user makes from the search engine *ad* to the marketed site. Even though these advertised content or SEM strategies are not usually effective [12], they are considered inefficient tools.

It was difficult for a neophyte business owner to push for a pie share from a small business perspective. With physical stores, they have none of the advantages big retailers get in business deals, often resulting in less competitive prices. Their opportunity resides in online selling. In this space, the predominance of big brands is not so colossal and there is space to them to boost their businesses.

It is undeniably that websites and e-commerce [13] can be customized in different optimizations brought by different technicians' expertise, paid advertised campaigns through their multichannel system or other SEM strategies push them as long as they can position the multiple before mentioned Search Engines.

These optimizations and strategies are also available to the small business owner, at least in the organic results. Without entering in technicalities, as we will do that later, we can present the general guidelines or good practices that Search Engines recommend:

- Use Keyword in titles and the content itself.
- Present unique content, description, and titles.
- Optimize Loading Speed.
- Use tools to track the site's performance, i.e., Google Analytics [14].
- Deliver content in next-gen formats.
- Provide internal links with the Web itself.
- Use Backlinks.
- Use Secure protocols i,e: HTTPS [15].
- Match queries with actual content.
- Present friendly URLs.

We can see a similar case in the "Black hat SEO " [16], consisting of a set of disapproved practices, that nevertheless, could end up pushing a Web's ranking, such as keyword stuffing, sneaky redirects, or hidden Text or Links. These practices will eventually get the website banned from Search Engine providers. The uncertainty of being de-indexed also makes this strategy frowned upon and risky. Moreover, we can also list a set of bad practices or things to avoid keeping a good SEO ranking:

- Keyword stuffing
- Low-quality guest posting for links
- Cloak text and links
- Present too many and cheap ads
- Too many outgoing links

Moreover, another advantage to the small owner/casual user is the trade-off between the cost of applying these strategies and the potential economic return. Therefore, the companies can obtain a substantial increase in visibility; consequently, the more visits a Web has, the higher probability to sell more products.

The evolution of the Web offered new possibilities. Nowadays, several tools allow a business to create a site. For instance, Wix [17] serves as an entry point to the casual user, who no longer fears the complex process of setting up and building a site.

The same trending platforms and other means to manage SEO also appeared over time. On the one hand, tools, like plugins, are managed manually by the user, like *YOAST SEO* [18], *The SEO Framework* [19], or *SEO PRESS* [20]. These free tools allow you to manage keywords presented in your pages, edit previews of results, scan well-known SEO bad practices, and redirection managing.

On the other hand, we have paid platforms. They do all of the above and more. Usually, the site owner is responsible for generating its content. Still, these platforms provide an *SEO friendly* range to the owner, manage social media in their place, and analyze the site owner's needs to generate an optimized Web structure. Considering this, more and more business owners seem to seek the opportunity to sell products and services online, using tools like newsletters, subscriptions models, and pay-per-view content to engage the audience.

Although SEO strategies apply to multiple Search Engines, it would be naive not concentrating on the dominant Search Engine, Google. Excepting some countries like China with the censored Search Engine Baidu [21] and Russia with Yandex [22]. So we will focus, mainly, on the scoring system presented by the very own tools Google presents to analyze Webs. In the second place, we will use impartial Webs or independent tools that track how well the SEO is constructed without targeting exclusive features in a single Search Engine.

With all the previous points in mind, this work will attempt to list, describe and analyze every possible action a Business owner with a site can do to upgrade his SEO score; What are these actions? How do they work? How much impact do they bring? or where to apply them?. Knowing our can or can not-do list, the owners of small businesses or SMEs will be able to apply the results of this work to better online visibility through the SEO algorithm optimizations we will be showing.

The main goal of this work is to list every modification on the site, highlighting its primary purpose regarding better positioning. This way, the author wants to show how to measure the impact of the different proposals. Let us consider the scenario where we compress all the images on the site as much as we can without an appreciable decline in quality. This modification will impact our loading time, which directly affects our numeric SEO Score.

The gist of our measures will be: Better HTML composition, image compression, deliver images in next-gen format [23] when possible, pre-load of assets like fonts, style sheets, and scripts files, eliminate render-blocking resources, Apply intelligent delivery of content through lazy load [24], include accessibility upgrades, like Accessible names for buttons, HTML names like iframes and images with a title and alt attribute [25], sizing the links correctly,... Try to assure security on our site, excluding cross-site links as much as possible and logging browser errors to the console [26].

2 State of the art

These are the different tools widely recognized and available to us to monitor SEO Results. We will be using the two first options, Lighthouse [27] a free tool and SEObility, a paid

one, to obtain the broadest possible information. These tools were used from 07/05/2021 to 07/18/2021.

- **Resource Optimization**: Meaning, checks for unused javascript, too large images, un-minified CSS [28], among other actions that may use computer resources (CPU, memory, or bandwidth).
- **Accessibility**: Various accessibility checks such as *alt* and *role attributes*, adequate font sizes, or ensure tap targets are big enough.
- **Link status**: Check from no broken links, canonical URLs [29], not too many cross-site links, among others.
- **Keywords**: Appropriate check for Keywords [30] and pages content (length, number of occurrences, or adequately positioned).
- **Meta Descriptions**: Ensure that every page has a meta description [31] with appropriate keywords and length.
- **HTML attributes**: There are a variety of HTML tags and attributes which ensure better SEO positioning, as *lang*, tag *DOCTYPE* [32], tag *viewport*, and others.
- **Image Serving**: Checks for next-gen images and correct formats.
- **Domain Checks**: Check for how the Web is placed in its domain [33].
- **Harmful tags**: Checks for some inadequate HTML tags you may use in your site.
- **Favicon**: Ensure that the site has a custom favicon [34], or one at all.
- **Server Configuration**: References multiple internal configurations that deliver a better SEO score (use HTTPS, SSL,...)
- **Mobile Optimization**: Offer suggestions on how to adapt your site to the mobile experience
- **Free**: Checks if it costs money to use this tool.

This research is not the first one of its kind, there is some previous groundwork we can see in sites like google scholar about how to SEO works and how to build efficient SEO mechanisms.

3 METHODS

This section presents and discusses the main items and categories used to upgrade our side regarding the metrics previously discussed. We present a list of different actions we can implement to enhance our SEO score.

- **Image compression**: Almost always, we will find images for our site in excellent quality, usually in JPG formats, or even worse, massive RAW [35] formats. However, these vast amounts of graphical information are underused when we show these images on our site in a reduced size, so the quality does not need to be the highest possible. Using tools like a JPEG compressor [36], we can significantly reduce the size of the image without it is discernible to the naked eye.
- **Setting limits**: There is a considerable range of different sizes, resolutions, and other properties from the same media content. For instance, the sites are commonly displayed at 4k [37] resolution, and usually, 1080p [38] screens. Furthermore, we will not be filling the entire screen with the image. So, it is useless to have images with massive resolution that increase loading times. Establishing default limits for media content, and sticking to them, will improve performance by reducing loading times.
- **Adopt new formats**: We have discussed RAW and `jpeg/jpg` formats, but these are old standards nowadays. The way to go is the `*.webp` [39] format, which significantly reduces the size of an image without losing quality. However, there are cons to this approach; older browsers like **Internet Explorer** do not support this new format.
- **Remove third-party media from your site**: Instead of using platforms like Vimeo [40] or Youtube [41], it is wiser to self-host your resources locally, in your server. Third-party software will need libraries and scripts that will bloat your site and affect loading times.
- **Edit video**: Using videos on your site can lower your SEO score by a considerable margin. The usual size of the videos will slow the connection and impede the interaction on the site for a long time. You can edit the video, lowering the resolution of the frame rate [42]. You can go even further and instead use a `GIF` [43].
- **Stop dissuading robots**: In order to place your site in the search engines, these must be able to crawl your site, telling them which pages to crawl and which ones should not be crawled. They expect a file called *robots.txt* [44]. Many platforms stop this behavior by default, trusting the user to change later when the Web is finished, or a user starts dissuading the robots while changing the site and never undoes his work.
- **Remove aria-hidden attributes**: Placing attributes aria-hidden in HTML tags will make assistive technologies like screen readers [45] skip the content; Google considers this a bad practice because they work inconsistently, and hidden content should be hidden for all users. So we will have to remove them and use other means like setting the CSS [46] *display* property to *none*.
- **Adding the role attribute**: Setting into our HTML footer tag with the value *contentinfo* identifies the section as information repeated at the end of every page, such as copyrights, navigation links, and privacy statements.
- **Proper Doctype in the HTML Document**: Instead of starting our HTML document with the tag <html>, we should declare on the first line a Doctype, in HTML5 it is done by adding <!DOCTYPE html>. It is information to the browsers telling him what kind of document to expect.
- **Use title tags**: In the HTML structure, we sure to add in the *head* [47] element a *title* element. This element tells the browser which text to show on your browser tab. This title will be displayed ad the link to click to go your site by the search engines.
- **Add a site description**: The site description is the text shown right below the page title to the user when

	Lighthouse	SEObility	WooRank	Sitechecker
Resources	✓	✓	✓	✓
Accessibility	✓	✗	✓	✓
Link status	✓	✓	✗	✗
Keywords	✗	✗	✓	✗
Meta Descriptions	✗	✓	✓	✓
HTML attributes	✓	✓	✓	✓
Image Serving	✓	✓	✗	✗
Domain checks	✗	✗	✓	✗
Harmful tags	✓	✓	✓	✓
Favicon	✗	✓	✓	✗
Server Configuration	✓	✓	✓	✓
Mobile Optimization	✓	✓	✗	✗
Free	✓	✗	✗	✗

Fig. 1. Comparison of different commercial software applications.

searching for a term in a search engine; a blank or wrong description will not offer the needed information to the user, so be sure to use an informative description with the appropriate term for your site.

- **Do not log JavaScript [48] errors to the console**: When loading scripts, it is usual to end up with errors on the browser console. Furthermore, we may be logging errors manually to the console. These facts will negatively affect us, so be sure to resolve each error or do a manual log.
- **Stop using JavaScript unload events**: These events are fired when an element is *unloaded or deleted* from the page. These events fire unreliably and cannot prevent browser optimizations. It is recommended to use page hidings or visibility change events.
- **Stop preventing default behavior on password inputs**: Some sites started preventing the user from pasting the password on the input field. This was seen as a bad practice, so now it is penalized to alter the default behavior on all inputs for passwords.
- **Remove AppCache references**: AppCache [49] was introduced in `HTML5` [50] and offered a *default* version of the Web for offline browsing. Nevertheless, this is already deprecated, and we should remove all references to it from our DOM [51] document and use the Cache API [52] instead.
- **Use the tag header**: Instead of using an `HTML` *div*element, use the specific *header* [53] tag, a tag for introductory content, and a set of navigation links.
- **Use the tag footer**: Instead of using an `HTML` *div* element, use the specific *footer* [54] tag, a tag for repeated content, and a set of navigation links and/or authorship information.
- **Use tag main, section, and article instead of divs**: Following the two previous items spirits, it is encouraged to preventively declare the content of an element by the tag of the content itself, and not just use div elements for everything. That means using the tag main for the main content of a document, the section tag for clearly defined and divided sections in the page, and the article tag for independent and self-contained content.
- **Add meta descriptions**: The meta description sketches the quality and content of a Web page.

This description is displayed for specific queries. The previous *site description* is also a meta description, but for the main page. Every page should have a meta description.

- **Add a language tag**: Right in the <**html** > tag, you should add a *lang* attribute, with the code ISO-639 [55] of your language as value. This action tells the screen readers to switch to another language.
- **Simplify HTML structure**: When using third-party tools to create your sites, it is usual to end up with DOM's consisting of thousands of nodes and child elements nested within each other. It has been recommended to have no more than 1500 nodes in total, no more than 32 levels of depth, and no more than 60 child nodes to a parent node.
- **Add a Favicon**: A favorite icon, also known as a shortcut icon or URL icon, is a file that contains small icons. These represent a brand or identity. Associating one with your site will override the shown icon in your browser's tabs.
- **Pages with at least 250 words**: Pages without at least 250 words on it will be penalized. It is an arbitrary number that used to be considered too much but nowadays is the expected quantity.
- **Add a viewport tag**: Add a viewport tag in the <**head**> tag, which sets the visible area o Web pages and instructs the browser on how to display them on different screen sizes (tablet/mobile/desktop).
- **Add *an alt* attribute to images**: The *alt* attribute tells the screen readers how to present an image by the value of the tag, usually a descriptive text of the image or the title of the section/product it shows. Empty or missing *alt* attributes are profoundly penalized, and it is a top priority to correct.
- **Use Keywords in your titles**: You should add relevant information in the form of the appropriate keywords you want to promote in your site in your title, especially in the more critical headers, such as <**h1**> tags.
- **Use a heading Hierarchy**: Heading tags outline sections and essential information, and they are used by assistive technologies to present an outline of the Web page, but they should follow an order and not skip intermediate tags. An <**h1**> will always precede an <**h2**>tag, and an <**h3**> will only have <**h3**> sibling in its hierarchy levels, and so on. Avoid using a smaller tag like an h5 or h6 after h2 or h1 tag
- **Minimize CSS**: The CSS file can be minimized by removing white spaces on it and performing other optimizations. Using tools like Prepros [56] for these operations will give you smaller files and faster loading times.
- **Divide CSS by pages**: A common mistake is to generate a single, massive CSS file and load it on every page. We can split the CSS files and smaller ones and load only the pertaining ones on each page, reducing the load on our server considerably.
- **Defer JavaScript files**: You can do so by adding the word defer to your <**script**> tag. When you defer [57] your JavaScript files, you load them asynchronously, parallel with the HTML load, without stopping the latter when the browser realizes you need a JavaScript file. Then it waits until the HTML is rendered and executes. This action guarantees that the files will be executed in the same order as they are loaded. This behavior is not guaranteed with the aysnc [58] loading of scripts.
- **Use Lazy Load for your images**: With lazy load is a pattern in which you do not load all the images at once. Only on-demand, for example, an image not visible on the screen will not be loaded until you scroll down and ends up visible. Lazy loading is almost enabled by default in third-parties software, or it gives the option to activate it. Furthermore, it is straightforward to find JS libraries to do it for you with ease.
- **Preload Fonts**: This tells the browser to start loading and caching a font immediately by using the rel attribute with the value *preload*, not when it is invoked in the HTML. The objective is to stop users from seeing a strange change of fonts and sizes while a page is loading, giving a bad user experience.
- **Remove permutations of fonts**: It is easy to load every permutation of font and not realize the strain it puts in the connection. For the font, you have the style of the font (italic, bold) and the weight (numeric value between 100 - 900). It is unnecessary to load every permutation of the font, instead select a few distinct ones, and the browser itself will render an approximate font.
- **Do not use unsafe cross-origin links**: When linking to another site on your page, be sure to avoid using the attribute *target* with value *blank*. This attribute indicates that the link should be opened in a new tab in the browser, not your current one. It is recommended to add a *rel* attribute with value **noopener** or **nofollow** which will avoid these issues.
- **Don't request the notification permission on page load**: Browsers have the capability of detecting your geolocation [59] using Javascript if you allow so. However, you should not ask for this permission right when the Web is loading, as it is considered it gives a bad user experience to get pop-ups just when you are loading the site. So stop requesting it, or only request it when it is going to be used.
- **Remove javascript vulnerabilities**: It is easy to install plugins and libraries, but many times, these get security concerns on themselves or their dependencies as time goes on. Inspect and remove older versions of vulnerable javascript and replace them with the new versions.
- **Use legible font sizes**: It is considered wrong concerning user experience to use small sizes for the fonts or try to hide texts to the screen readers by this attribute. Use the correct technique to fix it.
- **Size appropriately the tap targets**: Everything that can be tapped on your site in tablet and mobile resolutions should be big enough to tap. Otherwise, the user experience will be affected, and the search engine will penalize you accordingly.
- **Add a sitemap**: A sitemap helps with the organiza-

tion of your site, usually listing the URL and internal pages in a categorized way. This is done with a *sitemap.xml* [60] file hosted in a visible URL for the search engines.

- **Use HTTPS**: It is encouraged to offer your site with the HTTPS protocol, which uses an SSL certificate, helping secure or navigation and transactions and avoiding data stealing. Many payment platforms required and cutting edge web technologies require HTTPS to work on your site, so it's on our best interest to use it.

Every single item discussed here can be followed to this Github repository [1], where every item and rule was applied and measured in different commits (master branch).

4 Results

In this section, we will be presenting the results obtained with each tool, analyzing two pages on our site, the *Home page*, and the *Product page*, the Home Page refers to the front page of the site, the one that shows up at the root of its structure, the product page refers to the individual page every product has with details of the product itself and the buying options for that products . Every tool uses different titles for their metrics, so we will be showing the results using the specific name gathered from each tool on our demo page. We will be using Wordpress [61] to create this demo page with the url: https://clotheschest.udl.cat/, we decided to use wordpress for it's easy setup and ease of use.

First, let's explain the metrics used by the lighthouse tool. Performance refers to how fast a page loads and how fast a user can start interacting with it, the improvements for this score come from methods like reducing image size and the number of resources needed. Accessibility refers to how well our site can be accessed regardless of any user's impairments, and by how well the site maintains its functionalities for the most diverse range of users possible, one particular example would be the method of adding an alt attribute to image. Best practices refers to the overall code health of your site (make the page secure, create a good user experience and avoid deprecated technologies), the obvious example is the method of implementing HTTPS . Finally, SEO refers to a series of checks to ensure that the site is optimized for search engine results ranking, such as verifying how mobile friendly your page is, or if the site is available to robots and crawlers, also, making sure that search engines understand your content, here we can highlight the allowance of site crawlers method.

It is important to highlight that there will be slight dips in the results due to traffic. When we are measuring the performance traffic uncertainty may influence somehow the performance and the other scores. Furthermore, we need to keep in mind that our *performance* score will be capped by our hardware, i.e. The hosting server, as the primary sub metric for this one is how fast can the information be sent across the internet.

The results shown by Lighthouse tool are presented in Figures 2 (Home Page) and 3 (Product Page). These figures depict the relationship between the score obtained and the metric implemented on the web. Note that the x-axis represents the ID of a commit (1-41). Each commit belongs to the application of an improvement in the website; check GitHub history for a deeper explanation of each commit. This way, the y-axis represents the score obtained after applying the commit. Thus, these image assists in understanding the impact of each modification to the overall score regarding SEO.

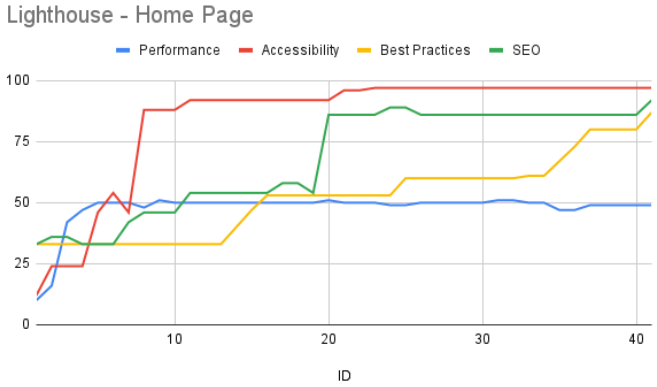

Fig. 2. Lighthouse Home Page Results

- **Performance**:The performance score in shown in our Figure 2 is affected by our images, the score rockets when we compress, reduce and use new formats for them (ID 1 - 3). Also, self-hosting our videos helps somewhat (ID 6).
- **Accessibility**: Having a speedy load helps with the accessibility score, so our image tampering helps (ID 26), but we other changes to help screen readers will do the real work here (removing hidden attributes, adding titles, descriptions, and language tags)
- **Best Practices**: The improvement in our score starts when we remove deprecated code written in JavaScript, and using the standard ways to login in to our site (Removing custom password inputs behaviors, ID 15). Using tags for the view port and removing obstructive notifications also helps. It is strongly encouraged to use HTTPS to obtain better scores as shown by the results in the Figure 2.
- **SEO**: Allowing robots to crawl our page (ID 6), adding titles and meta-descriptions raises the score, along some minor changes like using the correct viewport tag and again, using HTTPS (ID 41).

The results in our product page follow the same trend, but we can dig a little further on why the results behave like they do

- Performance: Our performance score here, as shown in the Figure 3, is always somewhat lower than in our Home Page. This happens because we serve a lot more images here. The image gallery display that goes along with the product information makes our servers struggle more.
- Accessibility: The results follow the same pattern as in our home page, but it is important to keep adding

1. https://github.com/imlmendez/clotheschest

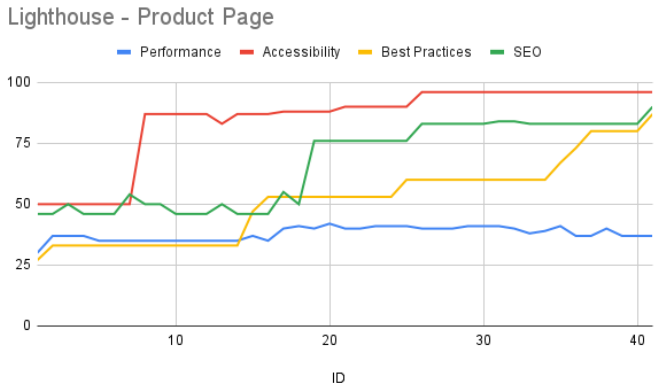

Fig. 3. Lighthouse Product Page Results

meta-descriptions and title to each and everyone of our products (ID 20, 10). Otherwise, Google will only position our Home page, but we also want to compete in the query's for sales.
- Best Practices: Again, we see the importance of HTTPS, without it, the user could be exposing himself to leak his personal information when buying, so it is even more important than in a normal site.
- SEO: When competing in the product query's, we can see how important it is to use our keywords correctly in our product titles, as to position ourselves above the competition (ID 27).

The results using the SEObility tool in our Home Page are the following:

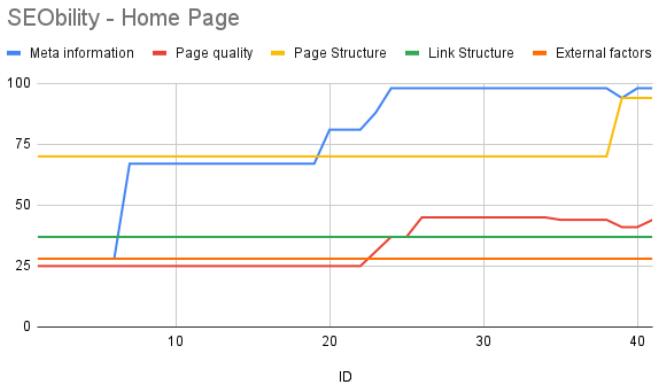

Fig. 4. SEObility Home Page Results

- Meta information: The name is self-explanatory: adding meta descriptions, favicons (ID 23), and allowing robots to crawl our site helps us with this metric.
- Page quality: A minimum number of words in the page (ID 24), favicons, and adding alternative texts to images helps improve this score significantly.
- Page structure: It refers, overall, to our heading structure, and how we use keywords in those headers.
- Link structure: This metric or bound by the technology or software providers we use, in some cases, we can choose how our URLs are generated, if so, it is important to avoid numbering of pages and excessive parameters in the URL.
- External factors: In this metric, we have little control over it. You should avoid serving adult content to keep out of blacklists, and invest in SEM so there are external links to the site across the internet. This is shown in the Figure 4 as a perfectly straight line in this metric.

And finally, these are the results with the SEObility tool for our product page.

- Meta information: Again, really important to use the title and description of the product to place the keywords we want to position (ID 11).
- Page quality: Using the appropiate HTML tags in header and footer helps us here, and keeping our gallery of images with alt description does too (ID 22).
- Page structure: The language tag and the keywords are once again the main protagonists for this metric.
- Link structure: The URLs for products should contain the title of the product itself as a keyword, and it should be a clear URL, without strange parameters in its structure.
- External factors: Idem as in out Home Page.

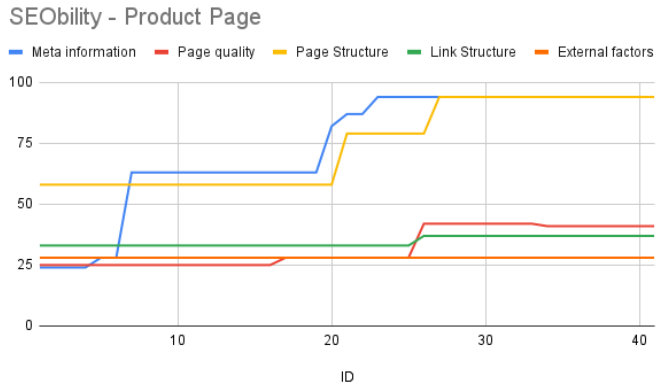

Fig. 5. SEObility Product Page Results

5 Conclusions

We can see a definite improvement in all the metrics observed for our site. So it is not far-fetched to qualify our results as encouraging. Some of the methods discussed and tested are automatically added by third party platforms, specifically those that delve in resource optimization, but we present a far more diverse and encompassing approach, optimizing every kind of metric we can. So, it is worth your to manually configure these kind of the techniques on your site, even more if you have a small business and need to compete as fiercely as you can. Improvements in these metrics will result in a better SEO score, and that translates to more visits and more product selling on your site, improving your revenue.

The tools used gave us good results, and although there is some overlap between the metrics they measure, we recommend using both if possible. But it must be said that the lighthouse one wins the race by a little margin in terms of usefulness, and how the metrics relate so closely to the google SEO ranking algorithm and being a completely free tool, as some features of SEObility are behind a paywall.

Once again, we present the methods needed to improve you site, but this time, the top 10 methods discovered while observing the results:

- Limit image sizes and compress them
- Use new formats for your images
- Allow robots to crawl your site
- Always use title tags
- Add meta descriptions for your products
- Don't be obtrusive when requesting permissions to the user.
- Use HTTPS
- Add Favicons
- add alt descriptions to images
- Remove/update old JavaScript libraries with vulnerabilities.

6 Future work

The online selling future resides in AMP (Accelerated Mobile Pages) [62], an open initiative launched by Google in 2015. We've seen how loading time affects SEO positioning, and AMP are all about mobile webs speedy loading times. As an open source coding standard for publishers it offers mobile speediness and responsiveness. It works using a customized subset of the HTML standard (amp HTML) []?, limiting the styles severely, only allowing "inline " [63] styling and less than 50 KB CSS, and lastly only allowing asynchronous Javascript. An AMP site is identified by a thunder icon on the google results page. This is done by adding special "amphtml " tag to your URL, telling google to render and execute the AMP page himself, usually with the help a pre-rendered CDN [64] server for cache. This is the key to the speediness we were talking about. We have to note, that it is technically difficult, and can present challenges to those who are less "tech-savy " and rely, like us, in frameworks (Wordpress) to ease in the online world. But it is a vital tool to the future SEO positioning ,and as such it is a necessity to the small business owner to not lag behind big enterprises in the ecommerce business.

References

[1] "A short history of the web — cern." [Online]. Available: https://home.cern/science/computing/birth-web/short-history-web

[2] "Starlink." [Online]. Available: https://www.starlink.com/faq

[3] N. Choudhury, "World wide web and its journey from web 1.0 to web 4.0." [Online]. Available: www.ijcsit.com

[4] "History of rss." [Online]. Available: http://www.rss-specifications.com/history-rss.htm

[5] "What is a search query? (definition) - seo glossary." [Online]. Available: https://growhackscale.com/glossary/search-queries

[6] "History and working of web crawlers — pdf — web search engine — search engine optimization." [Online]. Available: https://es.scribd.com/document/324297315/History-and-Working-of-Web-Crawlers-docx

[7] "10 amazon statistics you need to know in 2021 [march 2021]." [Online]. Available: https://www.oberlo.com/blog/amazon-statistics

[8] S. Chen and L. Lamberti, "Multichannel marketing: the operational construct and firms' motivation to adopt," *http://dx.doi.org/10.1080/0965254X.2016.1148759*, vol. 24, pp. 594–616, 11 2016. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/0965254X.2016.1148759

[9] "What is seo? search engine optimization 2021 - moz." [Online]. Available: https://moz.com/learn/seo/what-is-seo

[10] "What is sem? a guide to paid search engine marketing - alexa blog." [Online]. Available: https://blog.alexa.com/what-is-sem/

[11] "Google ads - aconseguiu més clients gràcies a una publicitat en línia fàcil." [Online]. Available: https://ads.google.com/intl/ca_es/home/

[12] "Why your sem strategy isn't working — cmg local solutions." [Online]. Available: https://www.cmglocalsolutions.com/blog/why-your-sem-strategy-isnt-working

[13] "e-commerce — definition, history, types, examples, facts — britannica." [Online]. Available: https://www.britannica.com/technology/e-commerce

[14] "Analytics." [Online]. Available: https://analytics.google.com/analytics/web/provision/#/provision

[15] "History of https usage." [Online]. Available: https://www.jefftk.com/p/history-of-https-usage

[16] "Ultimate black hat seo guide." [Online]. Available: https://seo-hacker.com/seo-black-hat/

[17] "About us — wix.com." [Online]. Available: https://www.wix.com/about/us

[18] "About us • yoast." [Online]. Available: https://yoast.com/about-us/

[19] "The seo framework • plugins that deliver results." [Online]. Available: https://theseoframework.com/

[20] "Seopress - best wordpress seo plugin." [Online]. Available: https://www.seopress.org/

[21] "Company overview — baidu inc." [Online]. Available: https://ir.baidu.com/company-overview/

[22] "Yandex." [Online]. Available: https://yandex.com/

[23] "Serve images in modern formats." [Online]. Available: https://web.dev/uses-webp-images/

[24] "What is lazy loading? - articles for developers building high performance systems." [Online]. Available: https://blog.stackpath.com/glossary-lazy-loading/

[25] "Image seo: How to optimize your alt text and title text • yoast." [Online]. Available: https://yoast.com/image-seo-alt-tag-and-title-tag-optimization/

[26] "Console overview - chrome developers." [Online]. Available: https://developer.chrome.com/docs/devtools/console/

[27] "Lighthouse — tools for web developers — google developers." [Online]. Available: https://developers.google.com/web/tools/lighthouse

[28] "What is minification — why minify js, html, css files — cdn guide — imperva." [Online]. Available: https://www.imperva.com/learn/performance/minification/

[29] "Consolidate duplicate urls with canonicals — google search central." [Online]. Available: https://developers.google.com/search/docs/advanced/crawling/consolidate-duplicate-urls

[30] "What are keywords? how to use them for seo." [Online]. Available: https://ahrefs.com/blog/what-are-keywords/

[31] "The importance of meta descriptions for seo — wordstream." [Online]. Available: https://www.wordstream.com/meta-description

[32] "Html doctype declaration." [Online]. Available: https://www.w3schools.com/tags/tag_doctype.asp

[33] "Beginner's guide: What is a domain name and how does it work?" [Online]. Available: https://www.wpbeginner.com/beginners-guide/beginners-guide-what-is-a-domain-name-and-how-do-domains-work/

[34] "Favicons - why they are important for seo." [Online]. Available: https://www.woorank.com/en/blog/the-importance-of-favicons-for-seo

[35] "What is raw in photography, and why should you shoot in raw? - finding the universe." [Online]. Available: https://www.findingtheuniverse.com/what-is-raw-in-photography/

[36] "What is a jpeg compressor?" [Online]. Available: https://www.easytechjunkie.com/what-is-a-jpeg-compressor.htm

[37] “4k resolution - wikipedia.” [Online]. Available: https://en.wikipedia.org/wiki/4K_resolution
[38] “What is 1080p resolution? fhd explained — tom’s hardware.” [Online]. Available: https://www.tomshardware.com/reviews/what-is-fhd-full-hd,5741.html
[39] “An image format for the web — webp — google developers.” [Online]. Available: https://developers.google.com/speed/webp
[40] “A history of vimeo, the site looking to usurp youtube — built in nyc.” [Online]. Available: https://www.builtinnyc.com/2016/05/02/brief-history-vimeo?__cf_chl_jschl_tk__=pmd_lC90TKvvkr5HZM7D31LSgM7Mk1pRSJqFmBNh7cj.n9A-1630528429-0-gqNtZGzNAfujcnBszQil
[41] “Youtube — history, founders, facts — britannica.” [Online]. Available: https://www.britannica.com/topic/YouTube
[42] “Blog menu.” [Online]. Available: https://www.techsmith.com/blog/frame-rate-beginners-guide/
[43] L. Nooney, L. Portwood-Stacer, and J. Eppink, “A brief history of the gif (so far),” *Journal of Visual Culture*, vol. 13, pp. 298–306, 2014. [Online]. Available: http://www.sagepub.co.uk/journalspermissions.nav
[44] “The importance of a robots.txt file for your seo.” [Online]. Available: https://seositecheckup.com/articles/the-importance-of-a-robotstxt-file-for-your-seo
[45] “A brief history of screen readers – knowbility.” [Online]. Available: https://knowbility.org/blog/2021/a-brief-history-of-screen-readers
[46] “A brief history of css until 2016.” [Online]. Available: https://www.w3.org/Style/CSS20/history.html
[47] “¡head¿: The document metadata (header) element - html: Hypertext markup language — mdn.” [Online]. Available: https://developer.mozilla.org/en-US/docs/Web/HTML/Element/head?retiredLocale=ca
[48] “Javascript history.” [Online]. Available: https://www.w3schools.com/js/js_history.asp
[49] “Appcache facts.” [Online]. Available: https://appcachefacts.info/
[50] “Html5: what is it? — techradar.” [Online]. Available: https://www.techradar.com/news/internet/web/html5-what-is-it-1047393
[51] “Javascript html dom.” [Online]. Available: https://www.w3schools.com/js/js_htmldom.asp
[52] “The cache api: A quick guide.” [Online]. Available: https://web.dev/cache-api-quick-guide/
[53] “¡header¿ - html: Hypertext markup language — mdn.” [Online]. Available: https://developer.mozilla.org/en-US/docs/Web/HTML/Element/header?retiredLocale=ca
[54] “¡footer¿ - html: Hypertext markup language — mdn.” [Online]. Available: https://developer.mozilla.org/en-US/docs/Web/HTML/Element/footer
[55] “Iso - iso 639 — language codes.” [Online]. Available: https://www.iso.org/iso-639-language-codes.html
[56] “Your friendly web development companion.” [Online]. Available: https://prepros.io/
[57] “Javascript defer - javatpoint.” [Online]. Available: https://www.javatpoint.com/javascript-defer
[58] “Html script async attribute.” [Online]. Available: https://www.w3schools.com/tags/att_script_async.asp
[59] “Geolocation api - web apis — mdn.” [Online]. Available: https://developer.mozilla.org/en-US/docs/Web/API/Geolocation_API
[60] “What is an xml sitemap and why should you have one? • yoast.” [Online]. Available: https://yoast.com/what-is-an-xml-sitemap-and-why-should-you-have-one/
[61] “About us.” [Online]. Available: https://wordpress.com/about/
[62] “What is amp? - the complete guide to accelerated mobile pages.” [Online]. Available: https://instapage.com/blog/amp
[63] “Inline styles in html — codecademy.” [Online]. Available: https://www.codecademy.com/articles/html-inline-styles
[64] “What is a cdn? — how do cdns work? — cloudflare.” [Online]. Available: https://www.cloudflare.com/es-es/learning/cdn/what-is-a-cdn/